Dynamics-based motion de-blurring for a PZT-driven, compliant camera orientation mechanism

Michael D. Kim Jun Ueda * †

Abstract

This paper proposes a method for removing motion blur from images captured by a fast-moving robot eye. Existing image techniques focused on recovering blurry images due to camera shake with long exposure time. In addition, previous studies relied solely on properties of the images or used external sensors to estimate a blur kernel, or point spread function (PSF). This paper focuses on estimating a latent image from the blur images taken by the robotic camera orientation system. A PZT-driven, compliant camera orientation system was employed to demonstrate the effectiveness of this approach. Discrete switching commands were given to the robotic system to create a rapid point-to-point motion while suppressing the vibration with a faster response. The blurry images were obtained when the robotic system created a rapid point-to-point motion, like human saccadic motion. This paper proposes a method for estimating the PSF in knowledge of system dynamics and input commands, resulting in a faster estimation. The proposed method was investigated under various motion conditions using the single degree-of-freedom camera orientation system to verify the effectiveness and was compared to other approaches quantitatively and qualitatively. The experiment results show that overall the performance metric of the proposed method was 27.77% better than conventional methods. The computation time of the proposed method was 50 times faster than that of conventional methods.

*M. D. Kim and J. Ueda are with the George Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA, USA. (mkim44@gatech.edu; jun.ueda@me.gatech.edu)

†This research was supported in part by NSF grants: Cyber-Physical Systems ECCS-0932208 and Dynamical Systems CMMI-1300019.

1 Introduction

The human eye is an organ that allows the brain to recognize the environment. Because the eye has a limited field of view (FOV), the human visual system scans the environment or changes the region of interest (ROI) by changing the direction of the person's gaze. The human eye has a total of three degrees of freedom (DOF), each generated by antagonistic pairs of extraocular muscles. Therefore, the human gaze is controlled by a total of six extraocular muscles, including four recti muscles and two oblique muscles [9]. It is known that human eyes have two representative movement patterns: smooth pursuit and saccade [12]. A saccade is one of the fastest and most accurate human movements [24]. Several physiology studies have shown that the movement time is within 50ms for 10 degree saccade or 50-80 ms for 30 degree saccade and the maximum velocity is 250 degree per second [17, 3, 4]. When saccades occur the human visual system receives blurry environmental information due to a finite integration time [39]. Evidence indicates that this blurred information, called motion smear, is partially compensated by neural processes [5, 8, 10] and that the human brain predicts or uses information on eye movements to reduce blurring [39, 32]. The suppression of motion smear also occurs during smooth pursuit [5]. From a control point of view, saccadic eye movements cannot be controlled by continuous visual feedback [20, 45]. Saccade occurs much more rapidly than proprioceptive, vestibular, or visual feedback can be returned to the brain [31]. This suggests that the saccade is completed in an open-loop manner [31]. Various studies have been conducted in the engineering field on camera positioning devices and image restoration techniques for blurry images caused by motion smear.

Various camera positioning devices inspired by the human ocular system have been designed. Several were designed to be driven by traditional servo motors with rigid links [23, 38, 26]. These mechanisms successfully generated a fast motion that is comparable to saccade but have little in common with biological muscle actuation. The human eye is oriented by means of antagonistic pairs of muscles that are driven by a single actuation system for each DOF. Also, heavy servo motors have a marked discrepancy from extraocular muscles. There are other notable positioning devices that are positioned by means of antagonistic pairs of alternative actuators such as piezoelectric or pneumatic actuators [42, 25]. However, these devices still have a discrepancy from muscle actuators because of their rigidity, an artifact of the use of traditional servo motors, in contrast with the contractility and compliance of extraocular muscles. Further, most of the mechanisms [23, 38, 42, 25] are controlled in a closed-loop manner, whereas physiological evidence indicates that saccade is completed in an open-loop manner [20, 45, 31]. Schultz *etal*. have developed a single DOF camera positioner oriented by antagonistic pairs of compliant cellular actuators that are controlled by an open-loop switching controller [34, 33, 35].

There are various approaches in the computer vision field to restoring blurry degraded images due to motion blur. Motion blur is caused by relative motion between an image acquisition system and an object during the exposure window. To restore an image that was degraded by motion blur, the motion path must to be estimated. The estimated path can be represented as a point spread function (PSF). The PSF is a blur kernel that describes the camera motion during an exposure window. If the motion blur is shift invariant, an un-blurred or latent image can be estimated with a deconvolution algorithm, given the estimated PSF and the obtained blurry image. Some studies proposed the estimation of the blur kernel from a single image [14, 18, 36], others from multiple images [37, 2]. Although these approaches successfully estimate the latent image, much computation time is required. In particular, the computation time greatly increases as the image size increases or as the blur kernel size increases. Raskar *etal*. proposed a fluttered shutter approach to controlling the exposure time, but this fluttering does not occur in the human vision system [29]. There are different approaches that do not analyze inherent properties of the blurry image but use external devices to estimate the motion blur. One study used an extra-high-speed camera with a low resolution to track the camera motion [6]. However, this approach also requires much computation time for post-image processing with the secondary camera. Another study used a gyro sensor and an accelerometer to track camera shake or motion [19]. However, it involves computations that estimate the path of a camera shake from noisy gyro and accelerometer sensors. Various approaches have been proposed to estimate the PSF for motion de-blurring, but such studies focused on blurry images that were manually taken from a high-end camera system, and not on the images obtained from the robotic system.

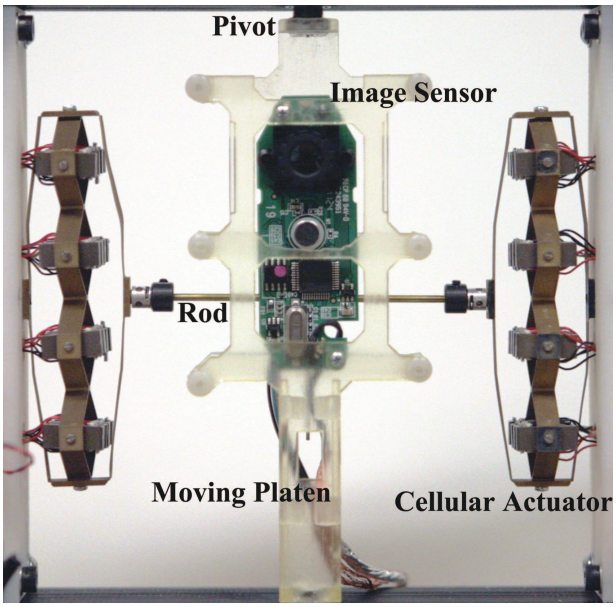


Figure 1: Biologically inspired camera orientation system [35]

All of the aforementioned studies on biologically inspired camera positioners were limited to mechanical design and control, without image processing. Also, image de-blurring studies have focused only on a hand-held camera but not on a robotic system. Previous studies focused solely on either mechanical design and control or image processing to estimate a latent image.

This paper proposes a study that connects the system

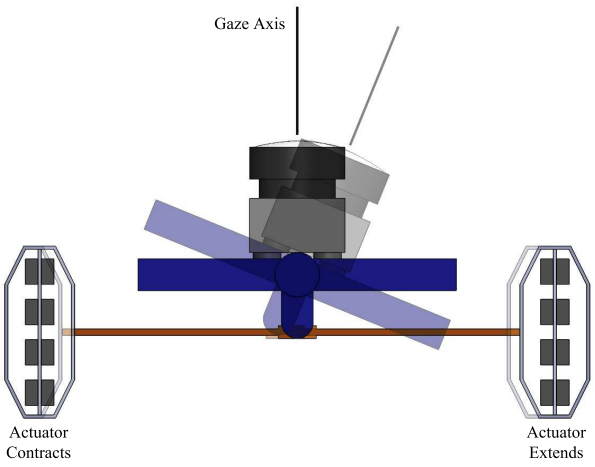


Figure 2: Pictorial representation of the camera orientation system. It is driven by an antagonistic pair of cellular actuators.

dynamics area and image processing area for a robotic camera positioner. Most of existing positioner systems use a high speed camera operating at around 100 frames per second (fps) while it is reported that the equivalent frame rate of the human eye is only around 24 fps. Therefore, if the system uses a camera with a frame rate similar to that of the human eye, the vision system capture blurry images during a saccade-like rapid motion. This biologically inspired system can aid the study in physiology to understand how the human brain processes image with ocular motor control. A robotic eye system can be employed with a high speed camera. However, such a system would still capture blurry images due to rapid motion even if the degree of blurriness is relatively low. Another situation that the high-speed camera may receive blurry images is low ambient light conditions.

This paper presents a dynamics-based estimation of the PSF. The study estimates the PSF via visual-motor coordination that is inspired by the observation that the human brain uses information on eye movements to reduce motion smear [39, 32]. The PSF can be directly estimated in an open-loop manner if the mechanical properties and control commands are known. Since the proposed method does not require additional sensors or analysis of inherent image properties, the computation time is much shorter for the estimation of a latent image. Although the proposed method can be implemented for a multi-DOF system, it was tested and verified using a single DOF vision system as shown in Fig. 1. This paper investigates the advantages and limitations of the proposed approach analytically and experimentally.

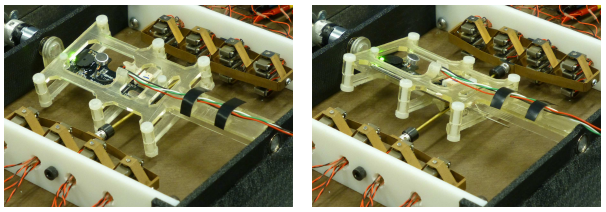
The proposed method requires an accurate calibration of the system dynamics. Otherwise, modeling errors could possibly reduce performance. The mechanical system in this study was modeled as a linear second-order system and any significant unmodeled dynamics was not observed. This paper focuses on fundamental issues associated with rapid camera orientation and image processing. It should be noted that the current system has some limitations: The vibration suppression technique is an open-loop approach and applicable only to a situation where disturbance is negligible or can be compensated by other means such as a vibration isolation mechanism. Future work will consider this issue.

2 The camera orientation system

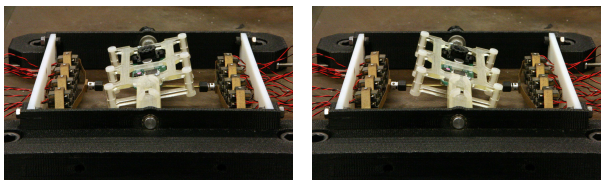
2.1 Mechanical Design

The camera orientation system used in this study is shown in Fig. 1. A Logitech C270 HD webcam is used for image acquisition. The plastic casings of the Logitech C270 were removed and only a board-level camera was installed on a moving platen. The rubber shielding of webcam USB cable was also removed to avoid introducing mechanical resistance to the camera orientation system. The moving platen is connected to a rod that transmits the force from antagonistic pairs of cellular actuators located on both sides. Each cellular actuator consists of sixteen Lead Zirconate Titanate (PZT) stack actuators with deformable amplification mechanisms [1, 40]. A PZT actuator has an extremely fast response and has high accuracy on the nano-meter scale with a low energy consumption. However, its small strain and small stroke distance limited the application to robotics. One approach to resolve the issue is to use a multilayered amplification structure. The multilayered strain amplification structure creates a larger displacement at the output point when forces are generated by the PZT stacks at the input. Since no gears or sliding mechanisms are used in the structures, this mechanism

exhibits zero backlash and noise-less operation while extremely fast movement is achieved. In addition, the nested compliant mechanism has been reported to resemble characteristics of human extraocular muscles [41]. The axis of the rod is positioned perpendicular but not orthogonal to the pivot axis. Therefore, when the rod is pushed or pulled by the antagonistic pairs of the cellular actuators, a moment is applied to the moving platen, resulting in a tilting motion of the camera positioning system as shown in Fig. 2. Figure 3 shows the operation range of the camera orientation system.



(a) Camera at extreme left (b) Camera at extreme right

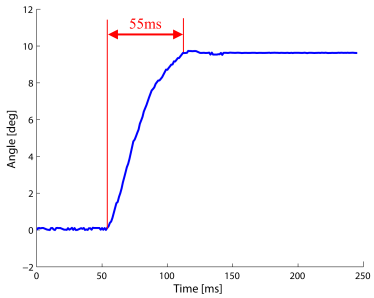


(c) Camera at extreme left (d) Camera at extreme right

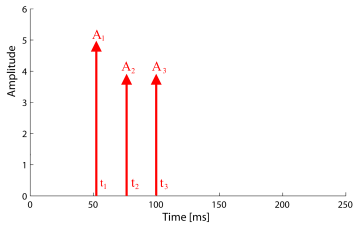
Figure 3: Single degree of freedom motion of the camera orientation system.

2.2 Motion Control

A total of 32 APA50XS (Cedrat corporation) PZT actuators are used to position the camera. The cellular actuator can be controlled continuously by adjusting the input voltage from 0V to 150V. The PZT actuator has a hysteresis response requiring additional sensors to estimate the position or to control the system. In this study, each PZT actuator is operated in an on-off manner to avoid hysteresis. This approach is similar to physiological bang-bang type of inputs given in human saccadic eye movement [13]. Since each cellular actuator has a total of 16 PZT actuators, the camera angle range is $\pm$ 16 PZT inputs or a tilting angle of $\pm$ 13 deg. The discretized desired angle



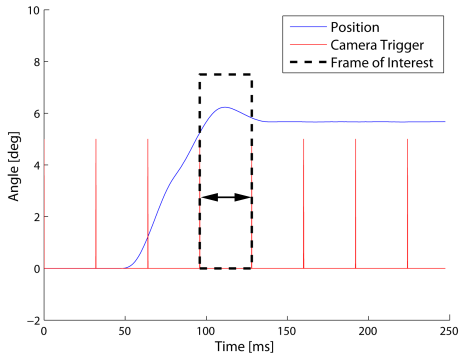
(a) Time response of a rapid point-to-point motion of the camera orientation system.



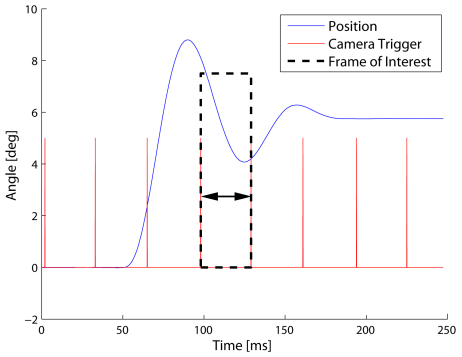
(b) Discrete switching commands to suppress vibration

Figure 4: A saccade-like motion with an open loop vibration suppression input

can be achieved by defining the numbers of activated PZT actuators. In this approach, the cellular actuators and constituent PZT actuators mimic human muscle physiology where whole muscle activity is a summation of motor unit activity, respectively [35]. Human saccadic eye motions are rapid eye movements within 50ms settling time at the maximum angular velocity of 250deg/sec. In order for the camera orientation system to generate the saccade-like motion, discrete switching commands must be given. The quantized discrete input commands can be determined by understanding the mechanical properties of the actuation system such as the natural frequency and the damping ratio. Figure 4(b) shows the discrete switching commands given to the system at appropriates time with appropriate amplitudes by applying a phase-vector analysis of mechanism control [34]. Figure 4(a) shows the time response of the camera orientation system as a result of the discretized switching commands. The results show that the vibration has been successfully suppressed and the settling time is



(a) Time response of the camera orientation device with vibration suppression input



(b) Time response of the camera orientation device with step input

Figure 5: Time responses of the camera orientation system.

within 55ms. If inappropriate commands or a step command are given to the system, however, the time response shows considerable oscillation resulting in a longer settling time.

2.3 Discrete Switching Input and Step Input

Figures 5(a) and 5(b) show the time responses of the camera orientation system given the vibration suppression input and the step input, respectively. When saccade takes place, the brain blocks images with substantial motion

(a) An image obtained with the discrete switching commands.

(b) An image obtained with the step command.

Figure 6: The image frame of interest obtained for both inputs, as indicated in Fig. 5.

blur in the course of eye motion which is called saccadic masking [27]. Therefore, the image frame of interest is the image obtained in the settling region. The image of interest from the time response with vibration suppression input is the third frame after the commands are given, as shown in Fig 5(a). For comparison, the same image frame was selected from the time response with step input, or vibration uncompensated input, as shown in Fig 5(b). The system starts obtaining images without motion blur after the fourth frame is given the vibration suppression commands, and the system starts obtaining images without motion blur after the sixth frame is given the step input. It can be easily seen that the camera orientation system experiences less vibration and has a faster settling time given the vibration suppression commands. It can also be seen from the obtained images that the system will receive less motion blur given the vibration suppression commands as shown in Fig 6.

2.4 Device Configuration

The discrete switching control is performed by a National Instruments (NI) cRIO-9024 embedded controller that is an Ethernet-based multi-core system. It runs VxWorks, Real-Time Operation System, at 1 kHz and FPGA at 40 MHz. Discrete switching commands are generated in the FPGA level because the switching delay is nominally

17μs. Voltage signals are given to the system by a NI 9401 5V DIO module with 100ns sampling time. The PZT is fully extended when 150V is given, so the voltage signals pass through a custom 5V to 150V amplification board. In Real-Time level, a NI 9401 module shares 5V/TTL signal with the host PC for syncing the image acquisition device and the mechanical control device.

The host PC receives images from the Logitech C270 at a fixed frame rate. The host PC is equipped with a NI PCI 6503, 5V/TTL DIO card, to sync with the embedded controller for visual-motor coordination. The testing environment is a PC running MS Windows 7 64bit version with Intel Core2 Quad CPU 2.83GHz, 4.00GB Ram and an ATI Radeon HD 4800 graphics card.

Software requirements must be addressed specifically for the dynamics-based motion de-blurring method. The main function is to synchronize the camera and the motion controller, which was achieved by sharing trigger signals. Poor synchronization would lead to misestimating the motion during the exposure window, which results in poor estimation of the PSF and the latent image. For the vibration suppression control, reasonable sampling rate of the motion control function is required due to a possible switching delay in commands. Otherwise, the switching commands would not be given to the system at appropriate times to suppress residual vibrations. The sampling rate of the motion control function was 1kHz. Given the information of control commands, the PSF was estimated in a parallel manner. After a blurry image was obtained, the deconvolution function was called to recover the latent image.

3 Image De-blurring

The blur kernel, or the PSF, represents a trajectory of the motion. In order to estimate the latent image accurately, the blur kernel must be estimated accurately. This paper proposes a novel method for the estimating of PSFs using the system dynamics of the camera orientation system.

3.1 Dynamics-based Blur Kernel Estimation

In the absence of external disturbances, the motion of a robotic camera orientation system is predictable since the dynamics of the system and input commands are known and given. Therefore, the motion of the system is estimated in an open-loop manner without the use of additional sensors. This approach is inspired by evidence that a saccade is completed in an open-loop manner [31] and the human brain predicts or uses information of eye movements to reduce motion smear [39, 32].

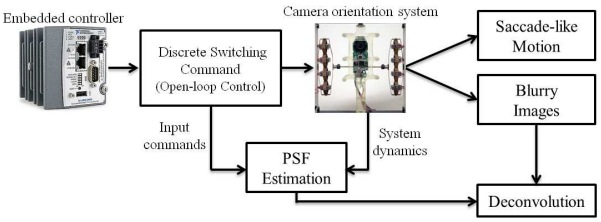


Figure 7: Dynamics-based image de-blurring for the robotic camera orientation system.

The state-space representation of a m-DOF, n-th order, linear system can be written as:

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \sum_{i=0}^{N-1} [\mathbf{B}\mathbf{u}_i(t-t_i)] \tag{1}$$

$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \sum_{i=0}^{N-1} [\mathbf{D}\mathbf{u}_i(t-t_i)] \tag{2}$$

$$\mathbf{u}_i(t-t_i) = A_i\delta(t-t_i) \quad (i \in 0,\cdots,N) \tag{3}$$

$$t_0 = 0 \tag{4}$$

$$\mathbf{x}(t) \in \mathbb{R}^{n\times 1}, \quad \mathbf{u}(t), \mathbf{y}(t) \in \mathbb{R}^{m\times 1}$$

$$\mathbf{A} \in \mathbb{R}^{n\times n}, \mathbf{B} \in \mathbb{R}^{n\times m}, \mathbf{C} \in \mathbb{R}^{m\times n}, \mathbf{D} \in \mathbb{R}^{m\times m}$$

where $\mathbf{x}(t)$ is the state vector, $\mathbf{u}(t)$ is the input vector, $\mathbf{y}(t)$ is the output vector, $\mathbf{A}$ is the system matrix, $\mathbf{B}$ is the input matrix, $\mathbf{C}$ is the ouput matrix, $\mathbf{D}$ is the feedthrough matrix, $\delta(t)$ is the unit inpulse function, A_i is the amplitude of i-th impulse, and N is the number of inputs given to the system. The linearity is assumed for simplicity in this paper, but nonlinear dynamics can also be used since the dynamics calculation is open-loop.

The camera orientation mechanism, employed in this study, exhibits a single dominant frequency at 14.4 Hz with a damping ratio of 0.19. However, the dominant frequency was different at 16.1Hz when the desired distance

was full stroke. The system parameters were identified experimentally by the observation of a step response. Therefore, the system can be represented as a linear second-order system given as:

$$G(s) = \frac{K}{s^2 + 2\zeta\omega_n s + \omega_n^2} \tag{5}$$

where K is the residue, ω_n is the natural frequency, and ζ is the damping coefficient. The value of K is 6650.3.

Since the robotic camera orientation mechanism has a single dominant natural frequency, it can be modeled as a linear, time-invariant (LTI), second-order, underdamped system. Then, the time response of the single DOF camera orientation system can be given as:

$$\mathbf{y}(t) = \theta(t)\hat{x} \tag{6}$$

$$\theta(t) = \sum_{i=1}^{j} \frac{A_i}{\omega_n^2}[1 - e^{-\zeta\omega_n(t-t_i)}\cos\omega_d(t-t_i) + \frac{\zeta}{\sqrt{1-\zeta^2}} \cdot \sin\omega_d(t-t_i))]u(t-t_i) \tag{7}$$

where A_i is the amplitude of the i-th step input, t_i is the time of the given i-th step input, j is the number of amplitudes given to the system, ω_n is the natural frequency, and ζ is the damping coefficient.

A PSF can be given as

$$k(x,y) = h_{(x,y)} \tag{8}$$

where h is an energy level at the location (x,y) of a pixel.

The pixel of the kernel (x,y) is

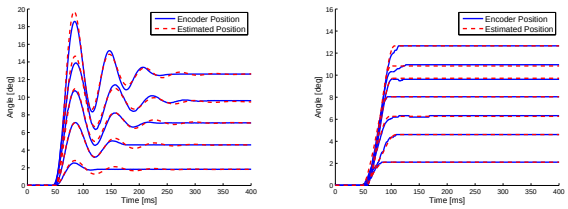
$$(x,y) = L\int_{t_a}^{t_b} \theta(t)dt\hat{x} \tag{9}$$

where L is a conversion factor dependent on the size of the blur kernel.

and the energy function is

$$h_{(x,y)} = \frac{t_b - t_a}{t_{acq}} \tag{10}$$

where t_{acq} is the acquisition time of a single image.



(a) Time response of the camera orientation system given the step command.

(b) Time response of the camera orientation system given the discrete switching commands.

Figure 8: Comparisons between the actual motion and the estimated motion.

Thus, the energy level at the location (x,y) of the PSF pixel is proportional to the time remaining at the location (x,y). The PSF is an energy distribution function for which the energy conservation constraint must hold:

$$\iint k(x,y)dxdy = 1 \tag{11}$$

Given the mechanical properties and input commands to the system, a discretized PSF can be estimated in an open-loop manner, without the use of external sensors as shown in Fig. 7. This approach results in the fast estimation of the PSFs using the robotic camera orientation system because it is estimated simultaneously with the image acquisition in an embedded motion controller.

The mechanical system was modeled as a linear second-order system. It has been reported by Schultz *etal*. that a linear second-order model is sufficient in the range of the motion of interest and any significant nonlinearity was not observed in the camera orientation mechanism when it was controlled by discrete switching commands [35]. Also, it has been reported by Schultz *etal*. that a linear dynamic model is sufficient for model-based control [34]. In addition, the experiment results comparing the real and calculated responses clarify the validity of the linear model as shown in Fig. 8. If interested readers would like to apply this control method to a general mechanism, they may need to resolve issues associated with possible nonlinearity in the structure.

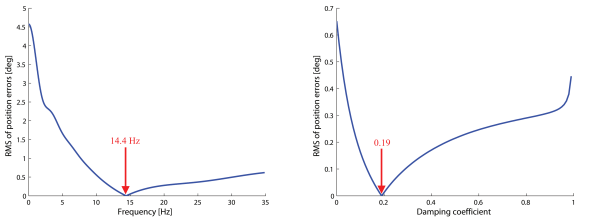
The proposed method requires an accurate estimation of the motion given the system dynamics and control commands. Figure 8 shows the experiment results of a comparison between the actual motion and the estimated mo-

tion. The actual motion was recorded by an encoder and the estimated motion was calculated by 7 implement on the FPGA. Figure 8(a) shows a comparison for step commands. For this particular mechanism, nonlinearity was not observed when the desired step was 10 degrees or less. A slight nonlinearity was observed where the natural frequency was 16.1 Hz when the desired step input was a full stoke step while the natural frequency was 14.4 Hz when the desired step input was below 77% of the full stroke. This was due to a saturation effect. The overshoot caused the third layer to collide with the first layer units [34]. Although the nonlinearity was observed, approximately a variation of 12% of the natural frequency, when the desired step was beyond 77% of the full stroke, the repeatability of the response was very high and the motion could be accurately estimated by using the predefined values of natural frequency accordingly as shown in 8(a). In this study, the range of step input where any significant nonlinearity was not observed was used as an input. The estimated motions show a good agreement with the actual motion overall with only minor errors after settling possibly due to flexibility of the cellular actuators and mechanical frictions.

Figure 8(b) shows the results for vibration suppression discrete switching commands where a good agreement is observed. When the discrete switching commands were used to generate a motion, any significant nonlinearity in the mechanism was not observed.

While the system in this paper was modeled as a linear second-order system, the dynamics-based motion deblurring method is theoretically applicable to a higher order linear dynamic system or a non-linear system. The motion of the system can be estimated by knowing the accurate dynamics model and control commands. However, the current method is not applicable to a system experiencing significant external disturbances.

This approach assumes that the system dynamics is modeled accurately and the motion has no unpredictable disturbances. Figure 9(a) and 9(b) show sensitivity analysis for system parameters given the discrete switching commands shown in Fig. 4(b). The sensitivity analysis was performed because the performance of the proposed method depends on the accuracy of system modeling. The system has no RMS position errors if the system parameters are calibrated accurately at frequency of 14.4Hz and damping coefficient of 0.19 as shown in Fig.



(a) Sensitivity plot for the frequency (b) Sensitivity plot for the damping coefficient

Figure 9: Sensitivity plots

9. The analysis shows that the motion is not estimated accurately with parameter errors, resulting in a poor estimation of the PSFs. In particular, the estimation model is not robust to the frequency parameter error. Although the performance is sensitive to the parameter calibration as expected, the calibration in general can be done with a reasonable precision. Any issues regarding the parameter calibration were not observed in this paper. It is possible to add command constraints on robustness to variations in frequency or damping coefficient but it results in more switching of commands and a longer settling time.

3.2 Deconvolution

A spatially-invariant blurred image $\boldsymbol{B}$ can be represented by a convolution between a shift-invariant blur kernel $\boldsymbol{K}$ and a latent sharp image $\boldsymbol{I}$ plus noise $\boldsymbol{N}$, as follows:

$$\boldsymbol{B} = \boldsymbol{K} \otimes \boldsymbol{I} + \boldsymbol{N} \tag{12}$$

$$\boldsymbol{B}, \boldsymbol{K}, \boldsymbol{I} \in \mathbb{R}^{m \times n}, \boldsymbol{K} \in \mathbb{R}^{l \times l}$$

where m is the width of the image in pixels, n is the height of the image in pixels, l is the length or height of the blur kernel, and $\otimes$ is the convolution operator.

As the PSF is obtained, the latent image can be estimated by means of a deconvolution algorithm. In this study, a widely-used Richardson-Lucy deconvolution method was selected [30]. This method is known to be robust in the presence of high noise levels. It estimates the latent image iteratively given the blurry input image and the estimated PSF [15].

$$\boldsymbol{I}^{n+1} = (\frac{\boldsymbol{B}}{\boldsymbol{I}^n \otimes \boldsymbol{K}} \otimes \hat{\boldsymbol{K}})\boldsymbol{I}^n \quad (13)$$

where $\boldsymbol{B}$ is the obtained blurry image, $\boldsymbol{K}$ is the blur kernel, $\hat{\boldsymbol{I}}$ is the flipped blur kernel, and $\boldsymbol{I}^n$ is the estimated image at n^{th} iteration.

For the deconvolution process, MATLAB's deconvlucy.m function was used.

3.3 Evaluation

In general, previous image de-blurring studies have qualitatively evaluated their results in comparison with those of other algorithms. Although the results of such evaluation were successfully discussed, it was a subjective evaluation. In this study, the results of the proposed de-blurring algorithm will be evaluated both qualitatively and quantitatively. The quantitative evaluation method subtracts the clean image with no motion blur from the estimated latent image and evaluates the pixel values of the difference image. The procedures of a quantitative evaluation method are as follows:

1. obtain a clean image with no motion blur at the closest discretized position to the blurry image;
2. convert both the estimated latent image and the clean image into grayscale;
3. pattern match between the estimated latent image and the clean image to find the translation matrix;
4. translate the clean image based on the previous step;
5. determine the difference between the estimated latent image and the translated clean image;
6. trim 10% of the image border to remove the boundary artifacts from the image deconvolution; and
7. evaluate

$$Difference\ Value = \frac{\sum_{j=1}^{n}\sum_{i=1}^{m} P_{ik}}{m \times n} \quad (14)$$

where m is the width of the image in pixels, n is the height of the image in pixels and P_{ik} is a pixel value at pixel (i,k).

The lower difference value results from the higher correlation between the estimated latent image and the clean image with no motion blur. Therefore, it can be concluded that the better de-blurring method will have a lower difference value.

4 Experiments

4.1 Compared Algorithms

Images with motion blur were obtained under various conditions to verify the proposed method. The results were compared with those of other notable and state-of the art algorithms proposed by: Fergus *etal*. [14], Shan *etal*. [36], Xu *etal*. [44], Goldstein *etal*. [16] and Whyte *etal*. [43].

Fergus *etal*. proposed a method to estimate the blur kernel by maximizing the probability in the image gradient domain [14]. Using the mixture of zero-mean Gaussians as the prior and given the measured image gradient, the posterior distribution can be defined. The method approximates the full posterior distribution and then estimates the blur kernel with maximum marginal probability, which is the maximum a-posteriori (MAP) solution. The approximation is made by a variation method, proposed by Miskin and MacKay [28]. The cost function, which is the gap between the approximating distribution and the true posterior, is minimized iteratively by the variation method. After obtaining the blur kernel, the Richard-Lucy algorithm is used to estimate the latent image [30].

Shan *etal*. found that the majority of the ringing artifacts in the latent image are caused by image noise and errors in kernel estimation [36]. Hence, they proposed a unified probabilistic model for both blur kernel estimation and latent image estimation to avoid the ringing artifacts. Both estimations are unified into a single MAP problem. Unlike Fergus *etal*. [14], the noise was modeled as a set of independent and identically distributed (i.i.d.) random variables with several constraints, which separates errors in image noise estimation and blur kernel estimation. The MAP problem minimizes the cost function, which is the negative logarithm of the proposed probability model. The cost function is optimized iteratively by alternatively updating the blur kernel and the latent image rather than using direct optimization because of its slow and poor convergence rate.

Xu *etal*. found that strong edges do not necessarily ben-

efit kernel estimation but deteriorate it under certain conditions. They proposed a two-phase blur kernel estimation to avoid the issue [44]. In the first phase, critical edges are defined by using Gaussian filtering and solving the shock filtering PDF problem. Then, the cost function, which is a function of noise, spatial prior, Gaussian regularization weight, is optimized to find a coarse blur kernel. In the second phase, a refined blur kernel is defined by iterative support detection (ISD) method to optimize a non-convex cost function. Finally, given the blur kernel, deconvolution was performed to find the latent image.

Goldstein *etal*. proposed a method to estimate the blur kernel from irregularities in the statistics of blurry images [16]. A power-law model was introduced to describe the power spectrum of natural images. After several processes such as filtering, approximation and whitening are performed, the 1D auto-correlation function of the differentiated projections of the blurry image is computed. Under non-negative energy, energy conservation, and finite exposure time assumptions, the full 2D power spectrum function of the blur kernel is approximated. Then, the blur kernel is iteratively estimated by a phase-retrieval algorithm given the minimized full 2D power spectrum function. Unlike the MAP-based methods [14, 36], the measured blurry images are not computed iteratively, resulting in cheap computation.

Whyte *etal*. proposed an estimation method of a non-uniform blur kernel that accounts for camera motion [43]. The non-uniform blur kernel includes a geometric model of camera motions: translation and rotation. Then, the algorithm proposed by Fergus *etal*. [14] was adopted and the algorithm proposed by Miskin and MacKay [28] was used to approximate the posterior distributions. The cost function is optimized by optimization of the Kullback-Leibler (KL) divergence between the posterior and the approximating distribution [7]. Since this method assumes that the rotational motion is a dominant factor in motion blur, the single DOF camera orientation system in this paper can be considered as a highly suitable application.

For a fair comparison, the parameters were adjusted a number of times to find the best possible blur kernels. The results were compared qualitatively and quantitatively.

4.2 Vibration Suppression Input and Step Input

Figure 12 shows the results of the proposed image deblurring method and others when the system was oriented by vibration suppression commands. The input blurry image is the frame of interest indicated in Fig. 5(a). The estimated PSFs are shown in Fig. 13. The intensity profile of the PSFs has a left-shifted center since the motion of the selected image frame was settled to the desired position. The amount of time that the device position was around the target position was dominant within a single exposure window, when the motion was created by the vibration suppression control as shown in Figure 5(a). Therefore, the intensity value in the PSF was accumulated relatively high around the target region, which corresponds to the left side of the PSF. The results show that all of the methods removed the blur because the vibration suppression technique reduced the motion blur. However, the results obtained by Goldstein *etal*., Shan *etal*., and Xu *etal*. present ringing artifacts. The results show that our proposed method, Whyte *etal*.s method, and Fergus *etal*.s methods have successfully estimated the latent image. The results can be compared quantitatively using the method suggested in section 3.3. The proposed method has the second best estimation according to the suggested quantitative evaluation.

Figure 14 shows the results of the proposed image deblurring method and others when the system was oriented by step input. The input blurry image is the frame of interest indicated in Fig. 5(b), which is the same frame in the previous case. The PSFs have a left-shifted center as shown in Fig. 15. This is because the motion of the selected image frame was oscillated around the second peak where the intensity is relatively high. However, the intensity profile is more distributed than the previous case because the velocity profile is almost constant except the oscillatory region at the end of the interested frame as shown in Fig. 5(b). The results show clearer comparisons than the previous case because the input image was more degraded. The results obtained by Goldstein *etal*., Fergus *etal*., and Whyte *etal*.s are still blurry. The results of Shan *etal*. and Xu *etal*. show an improvement, but the ringing artifacts still exist. The proposed method shows the most distinguishable letters and the best overall result. The proposed method also has the best quantitative result,

as shown in Fig. 25(b).

Comparing the results, it can be concluded that it is preferable to give the discrete switching commands as an input because it has a faster settling time and less motion blur in the image, resulting in better estimation of the latent image regardless of the image de-blurring method.

Figure 16 shows another experiment result when the system was capturing a difference scene and and oriented by vibration suppression commands. Figure 18 shows the estimated PSFs. Similarly to the first case, the intensity profile has a left-shifted center since the motion of the selected image frame was settled to the desired position. The results of Goldstein *etal*., Shan *etal*., and Xu *etal*. again present ringing artifacts and the results are slightly degraded. The methods proposed by Fergus *etal*. and Whyte *etal*. show an improvement but also have a small amount of ringing artifacts. The proposed method shows the best estimation with the least ringing artifacts among all the methods. The proposed method has the best quantitative result as shown in Fig. 25(c).

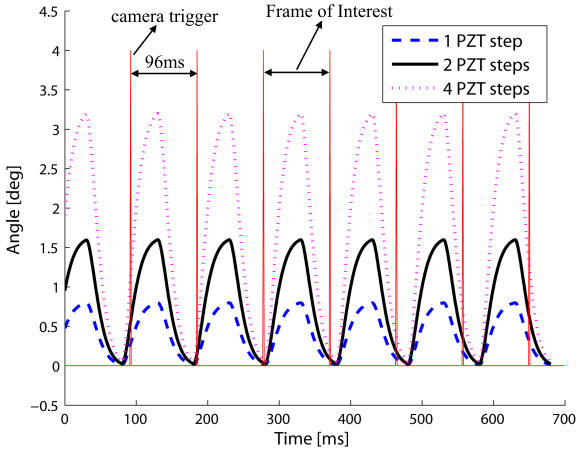


Figure 10: Oscillatory responses of the camera orientation system with various amplitudes. Images are obtained at 10.4fps.

4.3 Oscillatory Input

The proposed method has been discussed and compared to other approaches when the system experiences unidirectional movement. In this section, the proposed method was investigated when the system runs an oscillatory motion to verify that the proposed method performs effectively when the motion paths are overlapping or intersecting. For this experiment, it is preferable to have longer exposure times due to the limitation of working around the natural frequency of the camera orientation device. The longer exposure time ensures that the device experiences full-cycle of oscillation, but also make image acquisition susceptible to excess light. Therefore, images were obtained while indoor lights are off. Blurry images were obtained under oscillatory conditions with three different amplitudes as shown in Fig. 10. The amplitudes were one PZT step, two PZT steps, and four PZT steps. Images are obtained at 10.4 fps. The indicated image frame shown in Fig. 10 was investigated for all three amplitudes.

Figures 20, 22 and 24 show the estimated PSFs with different oscillation amplitudes. The PSFs have considerably uniform intensity distribution with low-amplitude multiple peaks. The amount of time staying at each pixel of the blur kernel within a single exposure window has a small variation as the camera was positioned. Therefore, the PSFs exhibit multiple peaks that is a natural consequence as the system experienced full cycle of oscillation.

Figure 19 shows image de-blurring results when the system obtained images while having an oscillatory motion created by one PZT step. Shan *etal*.s result is degraded due to an existence of dark noise in the capture image. Most of the methods show an improvement over the input blurry images but have small discrepancies from the image with no motion blur. The results of Goldstein *etal*. and Xu *etal*. have a small amount of ringing artifacts and the results of the proposed method and Whyte *etal*.s method are locally blurry. The proposed method, however, shows less blur globally.

Figure 21 shows image de-blurring results when the system obtained images while having an oscillatory motion created by two PZT steps. It can be checked again that Shan *etal*.'s method is not robust to low ambient light. The results of Goldstein *etal*. and Xu *etal*. are degraded with ringing artifacts and the result Fergus *etal*. still presents a large amount of blur. Although the results of the proposed method and Whyte *etal*. have a discrepancy with the image with no motion blur, those have the best results compared to the others.

Figure 23 shows image de-blurring results when the system obtained images while having an oscillatory motion created by four PZT steps. The results of Fergus *etal*. and Shan *etal*. have not estimated the blur kernels properly such that the letters are not distinguishable. The results of Goldstein *etal*. and Whyte *etal*. show an improvement but outputs are still blurry. Xu *etal*.' method has greatly improved the image but some of the characters are not clear. The proposed method shows the best qualitative result in that all characters except 'e' are recognizable. The proposed method has the second best quantitative result among the algorithms but the difference value is almost equal to that of the best result as shown in Fig. 25(f).

The effectiveness of the proposed algorithm was investigated under various motion conditions. The proposed algorithm estimated the PSFs overall better than conventional methods. The experiment results show that the performance metric of the proposed method is 12.4% better than the 2nd best algorithm and 27.77% better than the other algorithms on average. It can be concluded from the experimental results and the quantitative evaluations that the proposed algorithm has properly estimated the PSFs and showed better image de-blurring results than other approaches. However, residual blur or ringing exists in the image for some cases because it is theoretically impossible to recover high-frequency image from the images blurred with low-pass PSF.

4.4 Computation Time

Image de-blurring techniques is in general known to be computationally expensive. The proposed method, however, estimates the blur kernel in an open loop manner as the images are obtained. This approach results in a significant reduction of computation time. Figure 11 and Table 1 show computation times for each algorithm for each test presented in this paper. A total of 10 trials are conducted for each of the algorithm and for each test to evaluate average computation times. Computation times include estimation of a PSF and a deconvolution process. Computation times were measured by running either MATLAB or compiled executable programs distributed by the authors. For comparison, blur kernel sizes are fixed for each case. The properties that are used for comparing computation times are shown in Table 2. The results show that the proposed method is fastest among all the methods. This is because the PSF is estimated from the embedded motion controller simultaneously. The proposed method estimated latent images 4.8 times faster than the 2nd fastest algorithm and 51 times faster than the other algorithms on average. The computation time of the proposed method is not highly dependent on the size of the image and the blur kernel. On the contrary, the computation time is highly dependent on those for the methods proposed by Goldstein *etal*., Fergus *etal*., Whyte *etal*., and Shan *etal*. This is because those are based on a statistical model of the image or a series of optimization techniques. Most of the computation efforts of the proposed method result from the iterative deconvolution process. Cho *etal*. has shown slightly faster computation times in the literature [11] for some cases by executing a compiled C++ program in GPU. Since the proposed method was executed in MATLAB, the computation times can be reduced further by executing a compiled program in GPU or FPGA. In conclusion, the proposed method is the least computationally expensive.

Table 2: Image sizes that were used to estimate computation times

Properties	Image Size	PSF Size
Set #1	640x480	21x21
Set #2	640x480	21x21
Set #3	800x600	11x11
Set #4	640x480	11x11
Set #5	640x480	21x21
Set #6	640x480	35x35

Table 1: Comparisons of Computation Times [sec]

Methods	Proposed	Goldstein *etal*.	Fergus *etal*.	Whyte *etal*.	Shan *etal*.	Xu *etal*.
Software	MATLAB	MATLAB	MATLAB	MATLAB	Compiled Executable	Compiled Executable
Set #1	1.84	24.68	154.63	177.58	33.06	7.41
Set #2	1.17	24.76	144.66	168.39	32.88	7.41
Set #3	2.46	16.39	195.87	257.92	44.49	11.11
Set #4	1.12	15.47	148.19	162.37	20.45	5.31
Set #5	1.12	25.05	130.71	115.36	32.20	7.35
Set #6	1.88	47.59	111.50	172.45	140.16	7.61

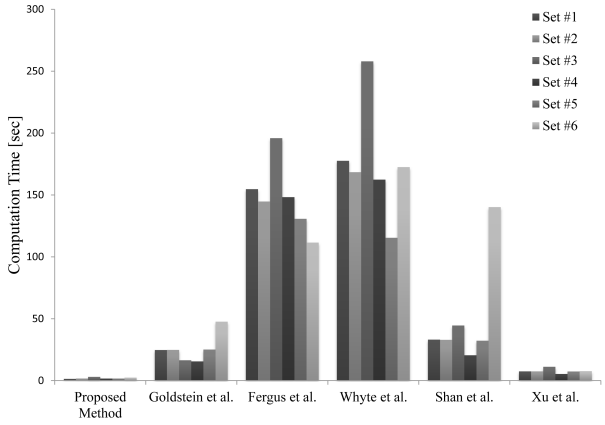


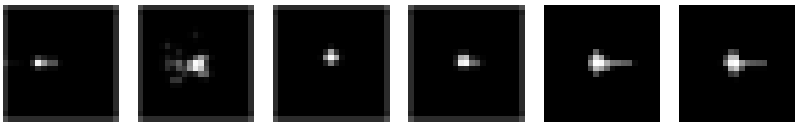
Figure 11: Comparisons of computation times

(a) Proposed Method (b) Goldstein *etal*. (c) Fergus *etal*.

(d) Whyte *etal*. (e) Shan *etal*. (f) Xu *etal*.

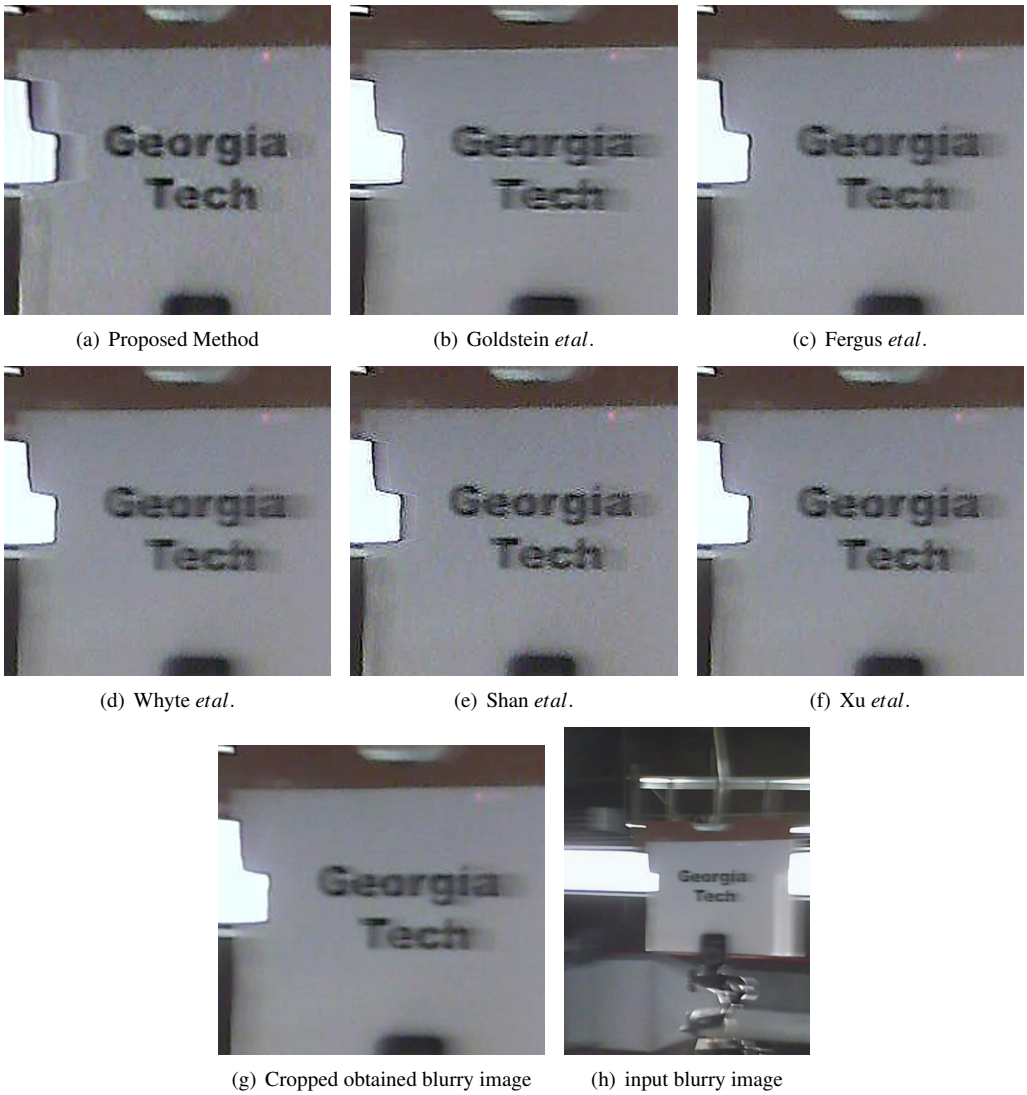
(g) Cropped obtained blurry image (h) Obtained blurry image

Figure 12: Test set #1. The obtained blurry image and estimated latent images by various methods. Images are taken at 29fps with Vibration Suppresion input command. (a) - (g) are cropped local images.



(a) (b) (c) (d) (e) (f)

Figure 13: Test set #1. Estimated blur kernels. The size of the PSFs is 21 $\times$ 21. (a) Proposed method (b) Goldstein *etal*. (c) Fergus *etal*. (d) Whyte *etal*. (e) Shan *etal*. (f) Xu *etal*.

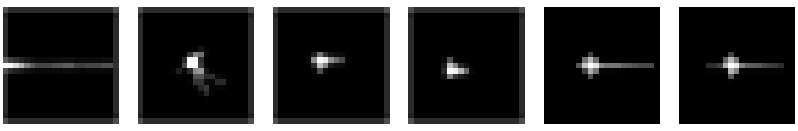


(a) Proposed Method (b) Goldstein *etal*. (c) Fergus *etal*.

(d) Whyte *etal*. (e) Shan *etal*. (f) Xu *etal*.

(g) Cropped obtained blurry image (h) input blurry image

Figure 14: Test set #2. The obtained blurry image and estimated latent images by various methods. Images are taken at 29fps with Step input. (a) - (g) are cropped local images.



(a) (b) (c) (d) (e) (f)

Figure 15: Test set #2. Estimated blur kernels. The size of the PSFs is 21 $\times$ 21. (a) Proposed method (b) Goldstein *etal*. (c) Fergus *etal*. (d) Whyte *etal*. (e) Shan *etal*. (f) Xu *etal*.

(a) Proposed Method

(b) Goldstein et al.

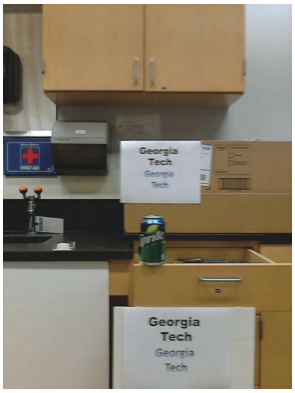
(c) Fergus et al.

(d) Whyte et al.

(e) Shan et al.

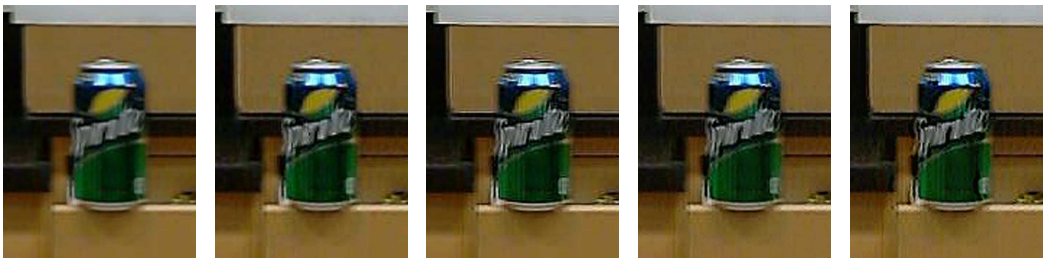
(f) Xu et al.

(g) Cropped obtained blurry image



(h) Obtained blurry image

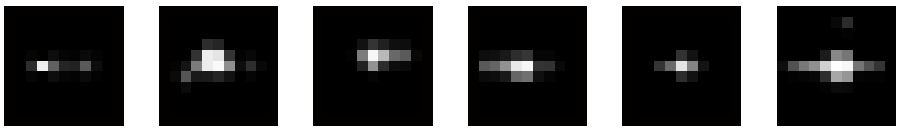
Figure 16: Test set #3. The obtained blurry image and estimated latent images by various methods. Images are taken at 29fps with Vibration Suppresion input command. (a) - (g) are cropped local images.



(a) Cropped obtained blurry image (b) Proposed Method (c) Goldstein et al. (d) Fergus et al. (e) Whyte et al.

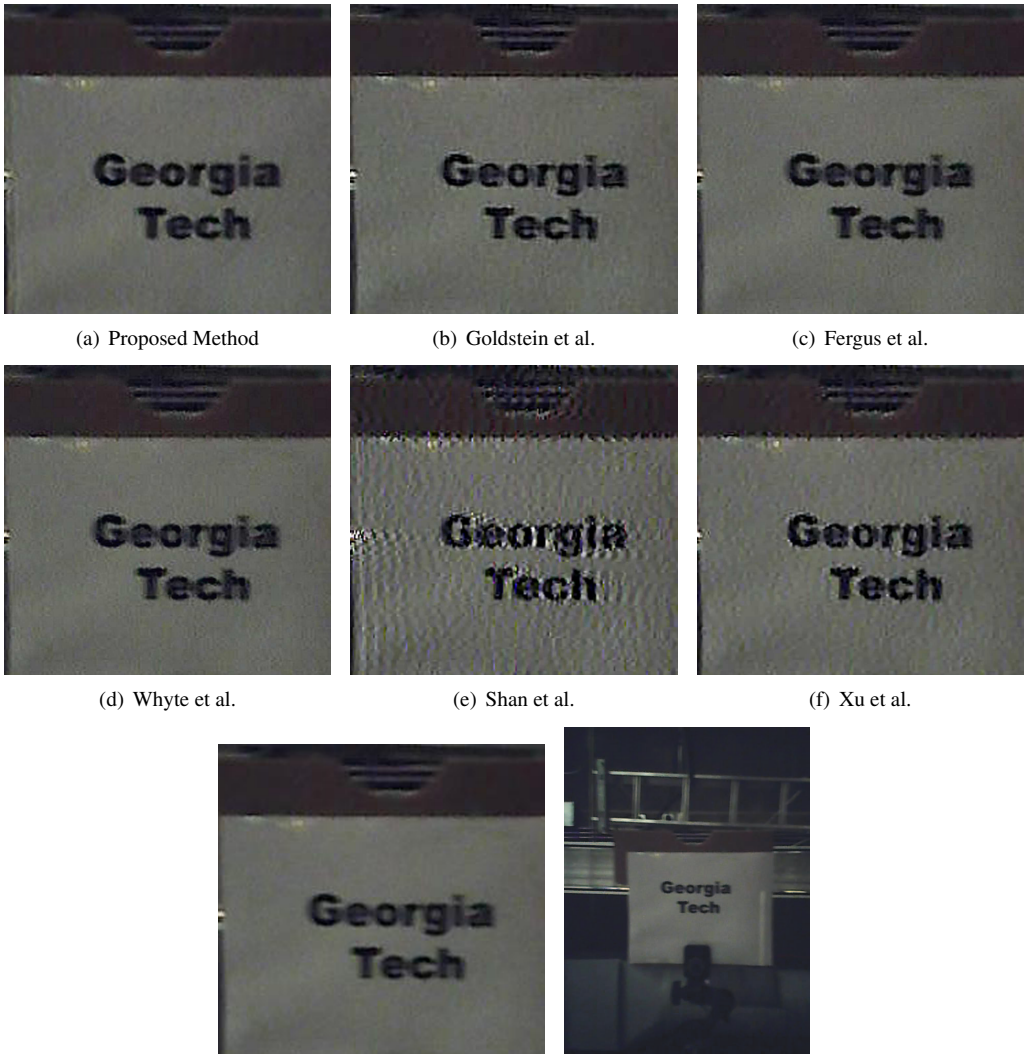
(f) Shan et al. (g) Xu et al.

Figure 17: Test set #3. The obtained blurry image and estimated latent images by various methods. Images are taken at 29fps with Vibration Suppresion input command. (a) - (g) are cropped local images.



(a) (b) (c) (d) (e) (f)

Figure 18: Test set #3. Estimated blur kernels. The size of the PSFs is 21 $\times$ 21. (a) Proposed method (b) Goldstein *etal*. (c) Fergus *etal*. (d) Whyte *etal*. (e) Shan *etal*. (f) Xu *etal*.

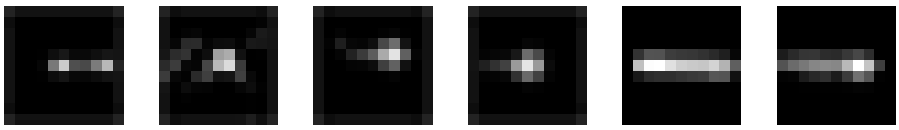


(a) Proposed Method (b) Goldstein et al. (c) Fergus et al.

(d) Whyte et al. (e) Shan et al. (f) Xu et al.

(g) Cropped input blurry image (h) Input blurry image

Figure 19: Test set #4. The obtained blurry image and estimated latent images by various methods. Images are taken at 10.4fps with periodic oscillatory input. One PZT step was used to create oscillatory response. (a) - (g) are cropped local images.



(a) (b) (c) (d) (e) (f)

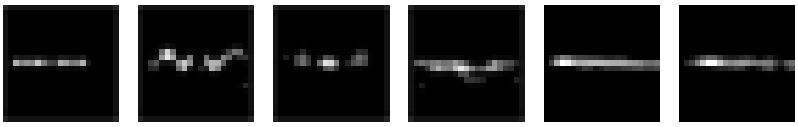
Figure 20: Test set #4. Estimated blur kernels. The size of the PSFs is 11 $\times$ 11. (a) Proposed method (b) Goldstein *etal*. (c) Fergus *etal*. (d) Whyte *etal*. (e) Shan *etal*. (f) Xu *etal*.

(a) Proposed Method (b) Goldstein et al. (c) Fergus et al.

(d) Whyte et al. (e) Shan et al. (f) Xu et al.

(g) Cropped input blurry image (h) Input blurry image

Figure 21: Test set #5. The obtained blurry image and estimated latent images by various methods. Images are taken at 10.4fps with periodic oscillatory input. Two PZT steps were used to create oscillatory response. (a) - (g) are cropped local images.



(a) (b) (c) (d) (e) (f)

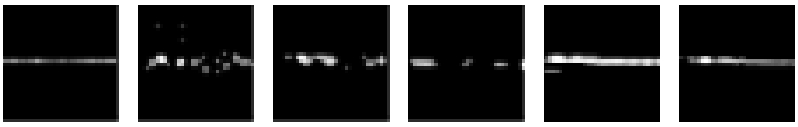
Figure 22: Test set #5. Estimated blur kernels. The size of the PSFs is 21 $\times$ 21. (a) Proposed method (b) Goldstein *etal*. (c) Fergus *etal*. (d) Whyte *etal*. (e) Shan *etal*. (f) Xu *etal*.

(a) Proposed Method (b) Goldstein *etal*. (c) Fergus *etal*.

(d) Whyte *etal*. (e) Shan *etal*. (f) Xu *etal*.

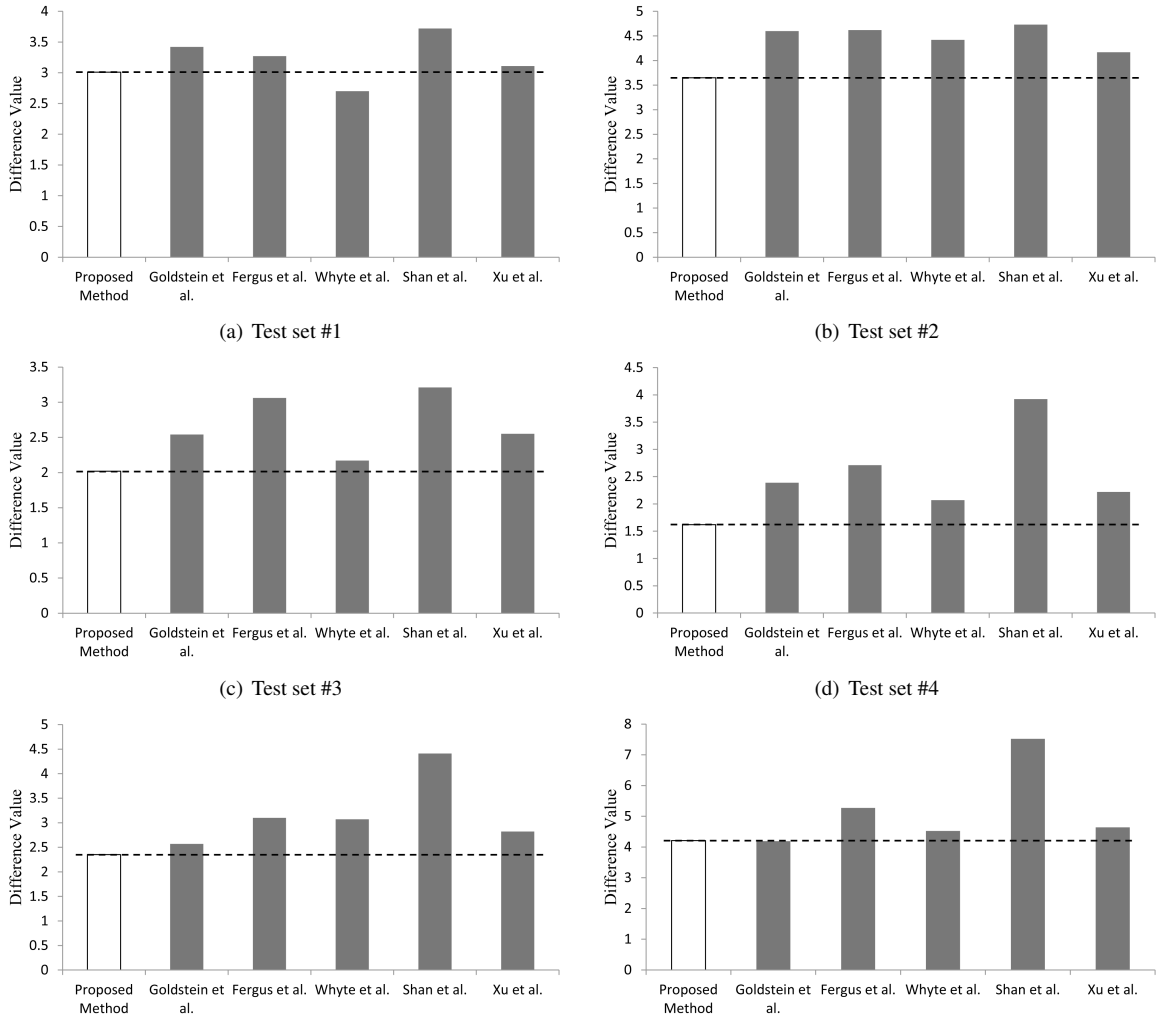
(g) Cropped input blurry image (h) Input blurry image

Figure 23: Test set #6. The obtained blurry image and estimated latent images by various methods. Images are taken at 10.4fps with periodic oscillatory input. Four PZT steps were used to create oscillatory response. (a) - (g) are cropped local images.



(a) (b) (c) (d) (e) (f)

Figure 24: Test set #6. Estimated blur kernels. The size of the PSFs is 35 $\times$ 35. (a) Proposed method (b) Goldstein *etal*. (c) Fergus *etal*. (d) Whyte *etal*. (e) Shan *etal*. (f) Xu *etal*.



(a) Test set #1

(b) Test set #2

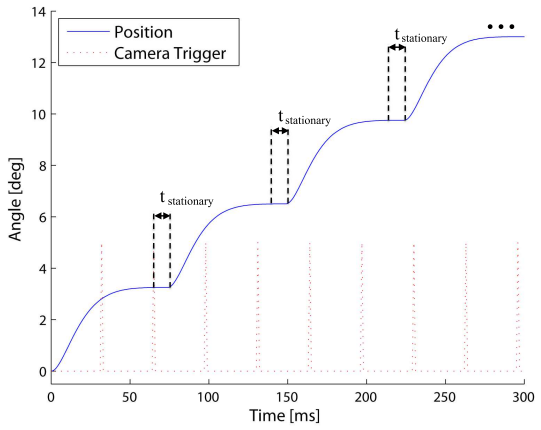
(c) Test set #3

(d) Test set #4

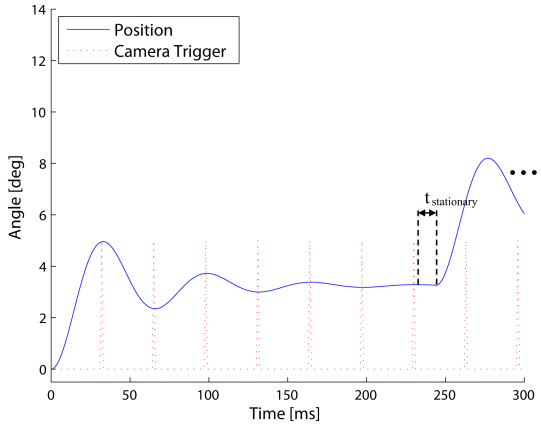
(e) Test set #5

(f) Test set #6

Figure 25: Quantitative evaluation results



(a) Visual-motor coordination with the vibration suppression input



(b) Visual-motor coordination with the step input

Figure 26: Visual-motor coordination

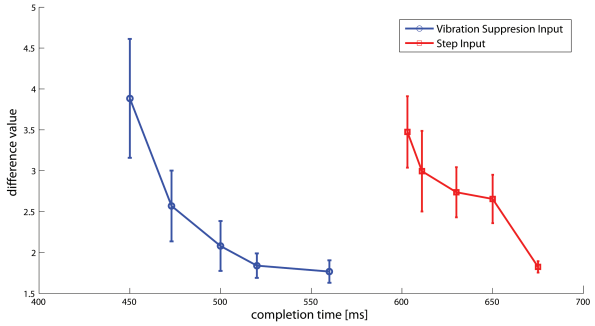


Figure 27: Results of visual-motor coordination. The camera orientation device scans four different positions. The total travel distance is 16 PZT inputs.

5 Visual-Motor Coordination

This section demonstrates the importance of the coordination between motion control and image processing for quick scanning of the field of view. While the proposed dynamics-based image processing technique is solely useful, as presented in the previous sections, the method is even more effective when combined with the vibration suppression technique in terms of scanning time. If the system needs to scan multiple positions, the total acquisition time is dependent on both mechanical settling time and $t_{stationary}$. $t_{stationary}$ is the difference between the time that the new commands are given and the time the motion generated by previous commands was settled in a single exposure window as shown in Fig 26. $t_{stationary}$ is minimally the exposure time of the camera if no motion was involved. The mechanical settling time is dependent on the natural frequency and can be considered as a constant. Therefore, the total acquisition time can be adjusted by $t_{stationary}$. A longer $t_{stationary}$ would result in less motion blur and vice versa. Since the degree of motion blur is dependent on the total acquisition time, there is a trade-off between a scanning time and an image quality.

Figure 27 shows the experimental results when the camera orientation device scans four different equally spaced positions. The proposed method was only used to estimate the latent image because its advantage has already been verified in Sec. 4. A total of four trials were conducted for each case and the total travel distance was 16 PZT steps. The results show that the completion times are faster overall when the vibration suppression commands are given. In comparison, the same quality of estimated latent images cannot be achieved in the same time by the step input, or vibration uncompensated input. When the vibration suppression commands are given, standard deviation of the estimated image quality decreases as the completion time increases. However, this tendency is not clear when the vibration uncompensated command is given. Therefore, the results show that

the vibration suppression technique is essential for the scanning time, even if the proposed de-blurring method is applicable for any type of commands. Also, quantitative evaluation supports the notion that the vibration suppression technique contributes to the image quality.

By means of visual-motor coordination, incorporating the dynamics-based motion de-blurring method with vibration suppression technique allows environmental scanning at a fast rate while achieving acceptable image quality.

6 Conclusion

This paper has presented a method to estimate the spatial-invariant PSF for the robotic camera orientation system. The robotic camera system was given discrete switching commands in an open-loop manner to effectively suppress the vibration. The camera orientation system captured images with relatively small motion blur as a result of the vibration suppression technique. The PSF was estimated in an open-loop manner with the aid of known system dynamics, requiring no external sensors. It was assumed that the system dynamics was well calibrated and the system had no unpredictable disturbances. The latent image was estimated by using the deconvolution process in knowledge of the estimated PSF.

The proposed method was investigated under various motion conditions to verify the effectiveness and was compared to well-known other approaches. The proposed method shows the best estimation of the PSF overall from the experiment results. The experiment results show that overall the performance metric of the proposed method is 27.77% better than that of conventional methods. In addition, the computation time of the proposed method is 50 times faster than the conventional methods. Although the proposed method was tested in the single DOF mechanism, it is also applicable to a multi DOF camera orientation system. Nonlinear dynamics can also be handled since the dynamics calculation is feedforward. The proposed method, however, is limited to the application that the camera is fully controlled by actuators.

By coordinating the motion and image processing of the camera positioner, the environment can be scanned at fast speed given the discrete switching commands. One can choose a point to scan the environment with a desired time while achieving acceptable image quality.

Future work includes an estimation of the spatial-variant PSF with a multi degree-of-freedom robotic camera orientation device. Also, the spatial-variant PSF may need to be estimated in knowledge of depth information from a stereo-camera system, when the robotic system is tracking an object. The proposed method has been tested by using a single degree-of-freedom system, but may be extended to a multi degree-of-freedom, stereo-vision system that is similar to the human ocular system.

7 Acknowledgment

The authors wish to express their thanks to Dr. Joshua Schultz for the valuable work on the camera positioning hardware, and to Ellenor Brown for the valuable discussions.

8 Appendix: Discrete Switching Commands

Since the robotic camera orientation mechanism is modeled as a second-order, LTI, underdamped system, the vibration can be suppressed by three impulses. The amplitudes and the times of three impulses can be determined from the vector diagrams in a complex plane as shown in Fig. 28 and solving the following equations:

$$\begin{aligned} A_0 \cos\phi_0 + A_1 \cos\phi_1 + A_2 \cos\phi_2 = 0 \\ A_0 \sin\phi_0 + A_1 \sin\phi_1 + A_2 \sin\phi_2 = 0 \end{aligned} \tag{15}$$

where ϕ_i is the direction of a phasor in the complex plane and A_i is the i-th impulse given to the system.

The direction of the phasor can be represented as:

$$\phi_i = \omega_n t_i \tag{16}$$

where ω_n is the natural frequency of the system and t_i is the time of the i-th impulse input given.

In order for the fastest solution, t_0 must be zero.

If the vector sum of three impulses is zero in the complex plane, the system will have no residual vibration after all three impulses are given. Since each PZT actuator is operated in an on-off manner, A_i must be an integer number.

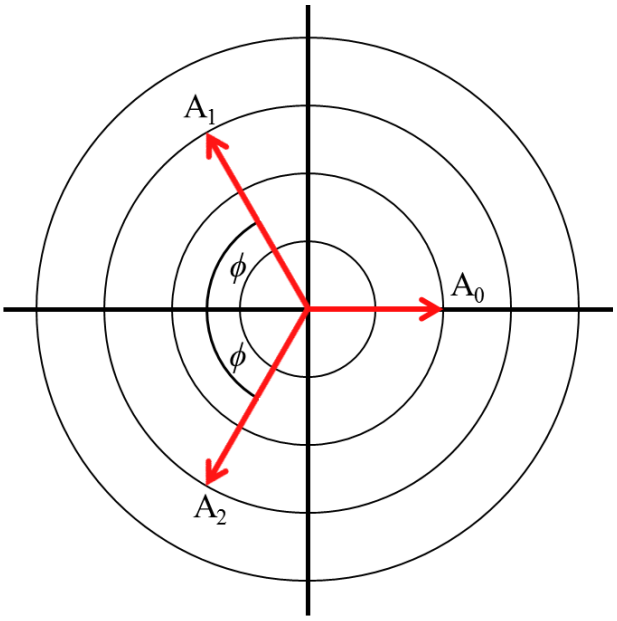


Figure 28: Representation of three impulses in the complex plane.

References

[1] Ehud Ahissar and Amos Arieli. Seeing via miniature eye movements: A dynamic hypothesis for vision. *Frontiers in Computational Neuroscience*, 6(89):1–27, 2012.

[2] L. Bar, B. Berkels, M. Rumpf, and G. Sapiro. A variational framework for simultaneous motion estimation and restoration of motion-blurred video. In *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*, pages 1–8, 2007.

[3] NH Barmack. Modification of eye movements by instantaneous changes in the velocity of visual targets. *Vision research*, 10(12):1431–1441, 1970.

[4] W. Becker and R. Jurgens. An analysis of the saccadic system by means of double step stimuli. *Vision Research*, 19(9):967 – 983, 1979.

[5] Harold E Bedell and Lori A Lott. Suppression of motion-produced smear during smooth pursuit eye movements. *Current Biology*, 6(8):1032–1034, 1996.

[6] M. Ben-Ezra and S.K. Nayar. Motion-based motion deblurring. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 26(6):689–698, 2004.

[7] Christopher M Bishop et al. *Pattern recognition and machine learning*, volume 1. springer New York, 2006.

[8] David C Burr. Motion smear. *Nature*, 284(13):164–165, 1980.

[9] Roger HS Carpenter. *Movements of the eyes (2nd rev)*. Pion Limited, 1988.

[10] Shuai Chen, Harold E Bedell, and Haluk Ö?men. A target in real motion appears blurred in the absence of other proximal moving targets. *Vision research*, 35(16):2315–2328, 1995.

[11] Sunghyun Cho and Seungyong Lee. Fast motion deblurring. *ACM Transactions on Graphics (TOG)*, 28(5):145:1–8, 2009.

[12] Raymond Dodge. Five types of eye movement in the horizontal meridian plane of the field of regard. *American Journal of Physiology–Legacy Content*, 8(4):307–329, 1903.

[13] John D. Enderle and James W. Wolfe. Time-optimal control of saccadic eye movements. *Biomedical Engineering, IEEE Transactions on*, BME-34(1):43–55, 1987.

[14] Rob Fergus, Barun Singh, Aaron Hertzmann, Sam T Roweis, and William T Freeman. Removing camera shake from a single photograph. *ACM Transactions on Graphics (TOG)*, 25(3):787–794, 2006.

[15] DA Fish, AM Brinicombe, ER Pike, and JG Walker. Blind deconvolution by means of the richardson–lucy algorithm. *JOSA A*, 12(1):58–65, 1995.

[16] Amit Goldstein and Raanan Fattal. Blur-kernel estimation from spectral irregularities. In *European Conference on Computer Vision (ECCV)*, pages 622–635. Springer, 2012.

[17] Christopher M Harris and Daniel M Wolpert. Signal-dependent noise determines motor planning. *Nature*, 394(6695):780–784, 1998.

[18] Jiaya Jia. Single image motion deblurring using transparency. In *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, pages 1–8. IEEE, 2007.

[19] Neel Joshi, Sing Bing Kang, C Lawrence Zitnick, and Richard Szeliski. Image deblurring using inertial measurement sensors. *ACM Transactions on Graphics (TOG)*, 29(4):30, 2010.

[20] R Jürgens, W Becker, and HH Kornhuber. Natural and drug-induced variations of velocity and duration of human saccadic eye movements: evidence for a control of the neural pulse generator by local feedback. *Biological cybernetics*, 39(2):87–96, 1981.

[21] M. D. Kim and J. Ueda. Dynamics-based motion deblurring for a biologically-inspired camera positioning mechanism. In *Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on*, page to appear, 2013.

[22] Falko Kuester and Tara C Hutchinson. A virtualized laboratory for earthquake engineering education. *Computer applications in engineering education*, 15(1):15–29, 2007.

[23] Yi-Chiao Lee, Chao-Chieh Lan, Cheng-Yu Chu, Chih-Ming Lai, and Yi-Jie Chen. A pan-tilt orienting mechanism with parallel axes of flexural actuation. *Mechatronics, IEEE/ASME Transactions on*, 18(3):1100–1112, 2013.

[24] RJ Leigh and Christopher Kennard. Using saccades as a research tool in the clinical neurosciences. *Brain*, 127(3):460–477, 2004.

[25] A. Lenz, S.R. Anderson, A.G. Pipe, C. Melhuish, P. Dean, and J. Porrill. Cerebellar-inspired adaptive control of a robot eye actuated by pneumatic artificial muscles. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 39(6):1420–1433, 2009.

[26] M. Lesmana and D.K. Pai. A biologically inspired controller for fast eye movements. In *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, pages 3670–3675, 2011.

[27] Ethel Matin. Saccadic suppression: a review and an analysis. *Psychological bulletin*, 81(12):899, 1974.

[28] James Miskin and David JC MacKay. Ensemble learning for blind image separation and deconvolution. In *Advances in independent component analysis*, pages 123–141. Springer, 2000.

[29] Ramesh Raskar, Amit Agrawal, and Jack Tumblin. Coded exposure photography: motion deblurring using fluttered shutter. *ACM Transactions on Graphics (TOG)*, 25(3):795–804, 2006.

[30] William Hadley Richardson. Bayesian-based iterative method of image restoration. *JOSA*, 62(1):55–59, 1972.

[31] D. A. Robinson. The mechanics of human saccadic eye movement. *The Journal of Physiology*, 174(2):245–264, 1964.

[32] Martin Rolfs, Donatas Jonikaitis, Heiner Deubel, and Patrick Cavanagh. Predictive remapping of attention across eye movements. *Nature neuroscience*, 14(2):252–256, 2010.

[33] J. Schultz and J. Ueda. A camera positioner driven by muscle-like actuation. In *Biomedical Robotics and Biomechatronics (BioRob), 2012 4th IEEE RAS EMBS International Conference on*, pages 719–724, 2012.

[34] J. Schultz and J. Ueda. Experimental verification of discrete switching vibration suppression. *Mechatronics, IEEE/ASME Transactions on*, 17(2):298–308, 2012.

[35] J. Schultz and J. Ueda. Nested piezoelectric cellular actuators for a biologically inspired camera positioning mechanism. *Robotics, IEEE Transactions on*, 29(5):1125–1138, Oct 2013.

[36] Qi Shan, Jiaya Jia, and Aseem Agarwala. High-quality motion deblurring from a single image. *ACM Transactions on Graphics (TOG)*, 27(3):73, 2008.

[37] Min Shi and Jiang Yu Zheng. A slit scanning depth of route panorama from stationary blur. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 1, pages 1047–1054 vol. 1, 2005.

[38] Yang Song and Xiaolin Zhang. An active binocular integrated system for intelligent robot vision. In *Intelligence and Security Informatics (ISI), 2012 IEEE International Conference on*, pages 48–53, 2012.

[39] Jianliang Tong, Scott B Stevenson, and Harold E Bedell. Signals of eye-muscle proprioception modulate perceived motion smear. *Journal of Vision*, 8(14):1–6, 2008.

[40] J. Ueda, T.W. Secord, and H.H. Asada. Large effective-strain piezoelectric actuators using nested cellular architecture with exponential strain amplification mechanisms. *Mechatronics, IEEE/ASME Transactions on*, 15(5):770–782, 2010.

[41] Jun Ueda. Piezoelectrically actuated robotic end-effector with strain amplification mechanisms. In *Advanced Mechatronics and MEMS Devices*, pages 25–52. Springer, 2013.

[42] T. Villgrattner and H. Ulbrich. Design and control of a compact high-dynamic camera-orientation system. *Mechatronics, IEEE/ASME Transactions on*, 16(2):221–231, 2011.

[43] Oliver Whyte, Josef Sivic, Andrew Zisserman, and Jean Ponce. Non-uniform deblurring for shaken images. *International journal of computer vision*, 98(2):168–186, 2012.

[44] Li Xu and Jiaya Jia. Two-phase kernel estimation for robust motion deblurring. In *Computer Vision–ECCV 2010*, pages 157–170. Springer, 2010.

[45] L.R. Young and L. Stark. Variable feedback experiments testing a sampled data model for eye tracking movements. *Human Factors in Electronics, IEEE Transactions on*, HFE-4(1):38–51, 1963.